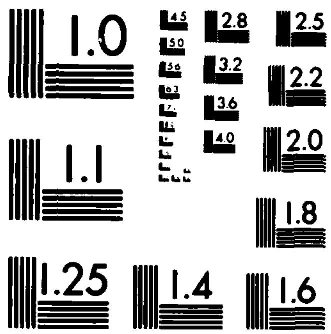
AD-A147 859 LOCAL DEPENDENCE AND POINT PROCESSES OF EXCEEDANCES IN STATIONARY SEQUENC..(U) NORTH CAROLINA UNIV AT CHAPEL HILL CENTER FOR STOCHASTIC PROC.. J HUSLER SEP 84 1/1

UNCLASSIFIED TR-77 AFOSR-TR-84-1054 F49620-82-C-0009 F/G 12/1 NL

END



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

CENTER FOR STOCHASTIC PROCESSES

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina

LOCAL DEPENDENCE AND POINT PROCESSES OF
EXCEEDANCES IN STATIONARY SEQUENCES

by

J. Hüsler

Technical Report #77

September 1984

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): TR-77	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 84-1054
6a. NAME OF PERFORMING ORGANIZATION: University of North Carolina; 6b. OFFICE SYMBOL (If applicable):	7a. NAME OF MONITORING ORGANIZATION: Air Force Office of Scientific Research
6c. ADDRESS (City, State and ZIP Code): Statistics Department, Phillips Hall 039-A, Chapel Hill NC 27514	7b. ADDRESS (City, State and ZIP Code): Directorate of Mathematical & Information Sciences, Bolling AFB DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR; 8b. OFFICE SYMBOL (If applicable): NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: F49620-82-C-0009
8c. ADDRESS (City, State and ZIP Code): Bolling AFB DC 20332-6448	10. SOURCE OF FUNDING NOS.: PROGRAM ELEMENT NO. 61102F; PROJECT NO. 2304; TASK NO. A5; WORK UNIT NO.

11. TITLE (Include Security Classification)
LOCAL DEPENDENCE AND POINT PROCESSES OF EXCEEDANCES IN STATIONARY SEQUENCES

12. PERSONAL AUTHOR(S)
J. Husler

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Technical	FROM ____ TO ____	SEP 84	15

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The point processes of high level exceedances are investigated under a weak mixing condition restricting the long range dependence of the stationary sequence. Depending on the local dependence, the asymptotic properties of the clustering of the exceedances and the cluster size distribution are found. This is described by the convergence of the point process of high level exceedances to a compount Poisson process.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
MAJ Brian W. Woodruff	(202) 767- 5027	NM

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

84 11 26 119

LOCAL DEPENDENCE AND POINT PROCESSES OF

EXCEEDANCES IN STATIONARY SEQUENCES

by

J. Hüsler

Technical Report #77

September 1984

LOCAL DEPENDENCE AND POINT PROCESSES OF EXCEEDANCES IN STATIONARY SEQUENCES

by

J. Hüsler

Summary

The point processes of high level exceedances are investigated under a weak mixing condition restricting the long range dependence of the stationary sequence. Depending on the local dependence, the asymptotic properties of the clustering of the exceedances and the cluster size distribution are found. This is described by the convergence of the point process of high level exceedances to a compound Poisson process.

1. Introduction

Let $X_1, X_2, \ldots$ be a stationary sequence of r.v, with $F(x)=P\{X_i \le x\}$. Assume that there exists a sequence $\{u_n\}_{n\ge 1}$ such that $n\cdot\overline{F}(u_n)\to\tau>0$, as $n\to\infty$, where $\overline{F}(x)=1-F(x)$. Furthermore we assume that Leadbetter's $D(u_n)$ holds, i.e. for any choice of integers

$$1\le i_1<i_2<\ldots<i_p<j_1<\ldots<j_{p'}\le n, \qquad j_1-i_p\ge \ell$$

for any n,ℓ,p,p' we have

$$|P\{X_{i_1}\le u_n,\ldots X_{i_p}\le u_n, X_{j_1}\le u_n,\ldots X_{j_{p'}}\le u_n\}-P\{X_{i_1}\le u_n,\ldots X_{i_p}\le u_n\}P\{X_{j_1}\le u_n,\ldots X_{j_{p'}}\le u_n\}|$$
$$\le \alpha_{n,\ell}$$

where for sequences $\{k_n\}$ and $\{\ell_n\}$

$$k_n\cdot\ell_n = o(n) \text{ and } k_n\cdot\alpha_{n,\ell_n} = o(1).$$

(see e.g. in Leadbetter, Lindgren and Rootzen (1983)).

For such a k_n define $r_n = [n/k_n]$ and let

$$N_n(i)=\sum_{j=(i-1)r_n+1}^{ir_n} 1(X_j>u_n) \qquad i=1,\ldots,k_n$$

be the number of exceedances in the ith block.

Similarly, let $N'_n(i)=\sum_{j=(i-1)r_n+1}^{ir_n-\ell_n} 1(X_j>u_n)$

be the number of exceedances in the ith block, where the last ℓ_n indices are deleted. In this case the blocks are separated by ℓ_n.

The cluster sizes $N_n(i)$ define now the marked point process Y_n on $(0,1]\times\mathbb{N}$ by setting: Y_n has a point at $(i/k_n, j)$ if $N_n(i) = j>0$. In the same way the marked point process Y'_n on $(0,1]\times\mathbb{N}$ is defined by replacing N_n by N'_n.

We give sufficient conditions on $\{X_i\}$ such that Y_n converges in distribution to a Poisson process Y on $(0,1] \times \mathbb{N}$. Define the projection π of point procysses in $(0,1] \times \mathbb{N}$ onto $(0,1]$ by setting $\pi(u) = \sum \beta_j \tau_{2j} \delta_{\tau_{1j}}$ for any point process $u = \sum \beta_j \delta_{\tau_j}$ with $\tau_j = (\tau_{1j}, \tau_{2j}) \in (0,1] \times \mathbb{N}$. Since the limit process Y is simple, $\beta_j \equiv 1$ and $Z = \pi(Y)$ is a compound Poisson process. By the continuous mapping theorem we find therefore that the point process $Z_n = \pi(Y_n)$ of exceedances of the level u_n in $(0,1]$ converges in distribution to Z, i.e. with the points i/r_n such that $N_n(i) \geq 1$.

The proof uses a theorem of Kallenberg (1976). Since the limit point process Y is simple, it is sufficient to prove that

(1) $\quad EY_n(B) \underset{n\to\infty}{\to} EY(B)$ for any $B = (a,b] \times N$ where $o<a<b\leq 1$ and $N \subset \mathbb{N}$

and

(2) $\quad P\{Y_n(B) = 0\} \underset{n\to\infty}{\to} P\{Y(B) = 0\}$ for any finite union B of disjoint "rectangles" as defined in (1).

In the second section we deal with the cluster size distribution, i.e. mainly with the statement (1). In Section 3 we give a sufficient condition such that (2) holds, which then implies our limit result. In the last section we discuss two examples, which exhibit two possible cluster size distributions which are concentrated in the first case on a finite number of points and in the other case on an infinite number of points.

2. The cluster size distribution

In this section we deal with the probability law of the marks on $\mathbb{N}$, i.e. with sufficient conditions such that

(A) $\quad P\{N_n(1) = k | N_n(1) \geq 1\} \underset{n\to\infty}{\to} \mu_k$ for all $k \geq 1$.

Define $E_n^{(s)} = k_n \sum_{1\leq i_1<i_2<\ldots<i_s\leq r_n} P\{X_{i_1} > u_n, \ldots, X_{i_s} > u_n\}$

for all $s \geq 1$, where u_n is such that $n\bar{F}(u_n) \to \tau > 0$

Let

$$s_o = \begin{cases} \inf \{s: \overline{\lim_{n\to\infty}} E_n^{(s)} = 0\} & \text{if such an s exists} \\ \infty \text{ otherwise} & \end{cases}$$

Then we use the following conditions:

(3) i) If $s_o < \infty$ assume that $E_n^{(s)} \underset{n\to\infty}{\to} \alpha_s$ for all $s \leq s_o$.

ii) If $s_o = \infty$ assume that $E_n^{(s)} \to \alpha_s$ for all s and that

$s^k \alpha_s \to 0$ as $s \to \infty$ for all $k < k_o$

where $k_o = \begin{cases} \inf \{k: \sum_{i=0}^{\infty} (-1)^i \binom{k-1+i}{k-1} \alpha_{k+i} = 0\} \text{ if such a k exists} \\ \infty \text{ otherwise} \end{cases}$

Furthermore we assume that

(4) $\quad \sum_{i=1}^{s_o} (-1)^{i-1} \alpha_i > 0$ where $E_n^{(s)} \underset{n\to\infty}{\to} \alpha_s$ for all $s \leq s_o$ and $\alpha_s \to 0$ if $s_o = \infty$.

This assumption is related to the case of a positive extremal index θ (see Leadbetter (1983) Corollary 3.5) since the sum in (4) is equal to $\theta\tau$. Thus we denote in the following $\sum(-1)^{i-1}\alpha_i = \theta\tau$.

We show now that an asymptotic measure μ on the marks exists, which gives the limit probability law in (A). By stationarity $P\{N_n(i) = k \mid N_n(i) \geq 1\} = P\{N_n(1) = k | N_n(1) \geq 1\}$ for all $i \leq k_n$. For the right hand side we have the following limit result.

<u>Lemma 2.1</u>: Assume that (3) and (4) hold. Then for all $i \le k_n$

$$P\{N_n(i) = k \mid N_n(i) \ge 1\} \to \mu_k \text{ for all } k \ge 1, \text{ as } n \to \infty,$$

where $\mu_k = \left(\sum_{i=k}^{s_0} (-1)^{i-k}\binom{i}{k}\alpha_i\right) / \left(\sum_{i=1}^{s_0} (-1)^{i-1}\alpha_i\right)$ for

$k < k_0$, $\mu_k = 0$ for $k \ge k_0$ with $k_0 = s_0$ if $s_0 < \infty$.

<u>Proof</u>. 1) We assume $s_0 < \infty$. By the Bonferroni-inequalities, for odd s_0,

$$k_n P\{N_n(1) \ge 1\} \le E_n^{(1)} - E_n^{(2)} + \ldots + E_n^{(s_0)}$$

and

$$k_n P\{N_n(1) \ge 1\} \ge E_n^{(1)} - E_n^{(2)} + \ldots - E_n^{(s_0-1)}$$

Taking the limit we find by using $\alpha_{s_0} = 0$ and (4)

$$\lim_{n\to\infty} k_n P\{N_n(1) \ge 1\} = \sum_{i=0}^{s_0} (-1)^{i-1}\alpha_i > 0,$$

which holds also if s_0 is even.

For $P\{N_n(1)\} \ge s_0\}$ we find that $k_n P\{N_n(1) \ge s_0\} \le E_n^{(s_0)} \to 0$ as $n \to \infty$;
thus $\mu_k = 0$ for all $k \ge s_0$.

Using the Bonferroni-inequalities again, we have for $s_0 - k$ even

$$k_n P\{N_n(1) = k\} \le E_n^{(k)} - \binom{k+1}{k} E_n^{(k+1)} + \ldots + \binom{s_0}{k} E_n^{s_0}$$

and

$$k_n P\{N_n(1) = k\} \ge E_n^{(k)} - \binom{k+1}{k} E_n^{(k+1)} + \ldots - \binom{s_0-1}{k} E_n^{(s_0-1)}$$

Taking the limit we have

$$\lim_{n\to\infty} k_n P\{N_n(1) = k\} = \sum_{i=k}^{s_0} (-1)^{i-k}\binom{i}{k}\alpha_i,$$

which is also true if $s_0 - k$ is odd. This implies our statement.

2) Let us assume now $s_0 = \infty$. As in 1) we have

$$\lim_{n\to\infty} k_n P\{N_n(1) \ge 1\} = \sum_{j=1}^{\infty} (-1)^{j-1}\alpha_j$$

since $\alpha_s \to 0$, and also

$$k_n P\{N_n(1) = k\} = \sum_{i=k}^{r_n} (-1)^{i-k}\binom{i}{k} E_n^{(i)}$$

Let $\bar{p}_k = \limsup_{n\to\infty} k_n P\{N_n(1) = k\}$, $\underline{p}_k = \liminf_{n\to\infty} k_n P\{N_n(1) = k\}$.

Then by the Bonferroni-inequalities we have for any even s and $k<k_o$

$$\sum_{i=k}^{s+k+1} (-1)^{i-k}\binom{i}{k}\alpha_i \le \underline{p}_k \le \bar{p}_k \le \sum_{i=k}^{s+k} (-1)^{i-k}\binom{i}{k}\alpha_i.$$

The two bounds differ by $\binom{k+s+1}{k}\alpha_{k+s+1} = O(s^k\alpha_{k+s+1}) = o(1)$ as $s\to\infty$, bv assumption (3). This implies as in the proof of Corollary 3.5 of Leadbetter (1983) that

$$\bar{p}_k = \underline{p}_k = \lim_{n\to\infty} k_n P\{N_n(1) = k\} = \sum_{i=k}^{\infty} (-1)^{i-k}\binom{i}{k}\alpha_i,$$

which gives our statement.

If $k_o<\infty$, we observe that $\mu_k = 0$ for all $k \ge k_o$, for

$$k_n P\{N_n(1) \ge k_o\} \le \sum_{j=0}^{j_o} (-1)^j\binom{k_o-1+j}{k_o-1} E_n^{(k_o+j)}$$

for any even j_o (see Feller (1968) p.110). Thus

$$\bar{p}_{k_o} \le \sum_{j=0}^{j_o} (-1)^j\binom{k_o-1+j}{k_o-1}\alpha_{k_o+j}$$

and similarly

$$\underline{p}_{k_o} \ge \sum_{j=0}^{j_o-1} (-1)^j\binom{k_o-1+j}{k_o-1}\alpha_{k_o+j}.$$

The two bounds differ by $O((k_o+j_o)^{k_o-1}\alpha_{k_o+j_o}) = o(1)$ as $j\to\infty$.

By the same argument as in i) we find

$$\bar{p}_{k_o} \le \sum_{j=0}^{\infty} (-1)^j\binom{k_o-1+j}{k_o-1}\alpha_{k_o+j} = 0. \qquad \square$$

<u>Remark 2.2</u> It is obvious that μ is a probability measure on the set of marks. For

$$\sum_{k=1}^{s_o}\sum_{i=k}^{s_o} (-1)^{i-k}\binom{i}{k}\alpha_i = \sum_{i=1}^{s_o}\alpha_i \sum_{k=1}^{i} (-1)^{i-k}\binom{i}{k} = \sum_{i=1}^{s_o}(-1)^{i-1}\alpha_i = \theta\tau.$$

To prove (1) we use weaker but less explicit conditions than (3) and (4). Suppose

$$k_n P\{N_n(1) \geq 1\} \to \theta\tau \tag{6}$$

for $\theta \in (0,1]$, $\tau>0$.

As mentioned above, (4) implies (6).

Lemma 2.3: If (A) and (6) hold, then

$$E(Y_n(B)) \to E(Y(B)) = \theta\tau(b-a)\mu(N) = \lambda \times \mu(B)$$

with B as in (1), $\mu(N) = \sum_{k\in N} \mu_k$ and $\lambda=\theta\tau m$ where m is Lebesque measure.

Proof: By stationarity we have

$$E(Y_n(B)) = \sum_{i\in(k_n a, k_n b]} P\{N_n(i)\in N\} \sim k_n(b-a)P\{N_n(1) \geq 1\}\ P\{N_n(1)\in N | N_n(1) \geq 1\}$$

$$\sim(\mu(N)+o(1))k_n(b-a)\theta\tau k_n^{-1} \to \theta\tau(b-a)\ \mu(N).$$

Remark 2.4: Obviously, condition (3) and (4) imply the assumptions of this lemma by Lemma 2.1; thus (1) holds if the explicit conditions (3) and (4) are satisfied.

Finally, we show that also the conditional probabilities in (A) with respect to N_n' converge to μ_k if (A) holds.

Lemma 2.5: Assume that (A) and (6) hold. If $k_n \ell_n = o(n)$, then

$$P\{N_n'(1) = k | N_n'(1) \geq 1\} \to \mu_k \text{ as } n\to\infty,\ k\geq 1.$$

Proof: We have for any $k\geq 1$

$$0 \leq P\{N_n(1) \geq k\} - P\{N_n'(1)\geq k\} \leq \sum_{j\in J} P\{X_j>u_n\} = \ell_n \bar{F}(u_n) = o(k_n^{-1}),$$

where $J = \{r_n-\ell_n + 1,\ldots,r_n\}$. Thus for $k \geq 1$

$$P\{N_n'(1) \geq k | N_n'(1) \geq 1\} = \frac{P\{N_n'(1) \geq k\}}{P\{N_n'(1) \geq 1\}} = \frac{P\{N_n(1) \geq k\} + o(k_n^{-1})}{P\{N_n(1) \geq 1\} + o(k_n^{-1})} = \frac{P\{N_n(1) \geq k\}}{P\{N_n(1) \geq 1\}} + o(1),$$

since $P\{N_n(1) \geq 1\} \sim \theta\tau k_n^{-1}$, $\theta>0$. This completes the proof.

3. The Poisson limit

In this section we mainly deal with the statement (2). In the first step we show that in (2) we may replace Y_n by Y'_n.

Lemma 3.1: If $k_n \cdot \ell_n = o(n)$, then for any B as in (2)

(7) $P\{Y_n(B) = 0\} - P\{Y'_n(B) = 0\} \to 0$ as $n \to \infty$.

Proof: Without any restriction we may assume that B is of the form

$$(8) \qquad B = \bigcup_{j=1}^{J} ((a_j, b_j] \times N_j)$$

where $0 \le a_1 < b_1 \le a_2 < b_2 \le \ldots \le a_J < b_J \le 1$, $J \ge 1$, $N_j \subset \mathbb{N}$.

For simplicity of notation we give the proof for $J = 1$, i.e. $B = (a,b] \times N$. Then the difference (7) is, with $I_n = (k_n a, k_n b]$, bounded by the two terms

$$P\{(N_n(i) \notin N, \forall i \in I_n) \cap (N'_n(i) \notin N, \; i \in I_n)^c\}$$

and

$$P\{(N_n(i) \notin N, \forall i \in I_n)^c \cap (N'_n(i) \notin N, \forall i \in I_n)\}.$$

The first term is bounded by

$$(9) \quad P\{\exists i \in I_n: \; N_n(i) \notin N \cap N'_n(i) \in N\} \le \sum_{i \in I_n} P\{N_n(i) \notin N \cap N'_n(i) \in N\}$$

$$= |I_n| \, P\{N_n(1) \notin N \cap N'_n(1) \in N\}$$

by stationarity. The last event implies that there exists some $j \in [r_n - \ell_n + 1, r_n]$ such that $X_j > u_n$. Thus (9) is bounded by

$$O(k_n \cdot \ell_n \cdot \overline{F}(u_n)) = o(1).$$

In the same way the second term is bounded. □

Thus it remains to prove that

$$(10) \qquad P\{Y'_n(B) = 0\} \to \exp\left(-\theta\tau \sum_{j=1}^{J} (b_j - a_j)\mu(N_j)\right)$$

where B is as in (8). The proof of this statement is simple, if the measure μ is concentrated in one point $k \in \mathbb{N}$. This is e.g.

the case if the condition $D'(u_n)$ of Leadbetter (see Leadbetter, Lindgren and Rootzén (1983)) holds; i.e. $\mu_1=1$, $\mu_k=0\ \forall\ k>1$. The proof follows the idea of the proof if $D'(u_n)$ holds.

Lemma 3.2: Suppose $D(u_n)$ and (6) hold. If (A) holds with $\mu_k=1$ for a $k\geq 1$ and $\mu_i=0\ \forall i \neq k$, then (10) is true.

In this particular case $\sum_{j=1}^{J}(b_j-a_j)\mu(N_j) = \sum_{j=1}^{J'}(b_j' - a_j')$, where we do not count the rectangles $((a_j,b_j] \times N_j)$ with $N_j \cap \{k\} = \phi$, i.e.

$$B \cap ((0,1] \times \{k\}) = \bigcup_{j=1}^{J'} ((a_j',b_j'] \times \{k\}) = B'.$$

Proof: i) We show that $P\{Y_n'(B) = 0\} = P\{Y_n'(B') = 0\} + o(1)$. Since $Y_n'(B) \geq Y_n'(B')$ we have for $I_n = \bigcup_{j=1}^{J} (k_n a_j,\ k_n b_j]$

$$0 \leq P\{Y_n'(B') = 0\} - P\{Y_n'(B) = 0\} \leq \sum_{i\in I_n} P\{N_n(i) \in \mathbb{N} \setminus \{k\}\}$$

$$= |I_n| P\{N_n(1) \geq 1\} \cdot P\{N_n(1) \neq k \mid N_n(1) \geq 1\} = O(k_n \cdot k_n^{-1}\ o(1)) = o(1),$$

since (6) implies $P\{N_n(1) \geq 1\} \sim \theta\tau/k_n$.

ii) Let $I_n' = \bigcup_{j=1}^{J'} (k_n a_j', k_n b_j']$. Since $\{N_n'(i) = 0,\ \forall i \in I_n'\}$ $\subseteq \{N_n'(i) \neq k,\ \forall i \in I_n'\} = \{Y_n'(B') = 0\}$, it follows that

$$0 \leq P\{N_n'(i) \neq k,\ \forall i \in I_n'\} - P\{N_n'(i) = 0,\ \forall i \in I_n'\} \leq \sum_{i\in I_n'} P\{N_n'(i) \notin \{0,k\}\}$$

$= O(k_n \cdot k_n^{-1} o(1)) = o(1)$ as in i).

iii) Let $A_i = \{N_n'(i) = 0\}$. Then by enumerating the i's in I_n' as $i_1, i_2, \ldots, i_{j_n}$ with $j_n = |I_n'|$, we get

$$\left|P\{N_n'(i) = 0,\ i \in I_n'\} - \prod_{i\in I_n'} P\{N_n'(i) = 0\}\right| \leq$$

$$(10)\quad \le \sum_{j=1}^{j_n-1} \left| P\left(\bigcap_{k=1}^{j+1} A_{i_k}\right) - P\left(\bigcap_{k=1}^{j} A_{i_k}\right)\cdot P(A_{i_{j+1}})\right|.$$

Since the index sets are separated by ℓ_n, it follows by $D(u_n)$ that each term in (10) is bounded by α_{n,ℓ_n}. Since $|I_n'| = j_n = O(k_n)$ we see that (10) is bounded by $O(k_n\cdot\alpha_{n,\ell_n}) = o(1)$ by $D(u_n)$. Finally

$$\prod_{k=1}^{j_n} P(A_{i_k}) = [P(A_1)]^{j_n} = \left(1-\frac{\theta\tau + o(1)}{k_n}\right)^{j_n} \to \exp(-\theta\tau\sum_{j=1}^{J'}(b_j' - a_j')) \text{ as}$$

$n \to \infty$, by (6). □

We have therefore in some particular cases the desired result.

Theorem 3.3: Assume that the stationary sequence $\{X_k\}$ satisfies $D(u_n)$, (A) and (6). If the measure μ is concentrated in a single point $k \geq 1$, then

$$Y_n \xrightarrow{d} Y \text{ as } n \to \infty$$

where Y is a Poisson process on $(0,1] \times \mathbb{N}$ (concentrated on $(0,1] \times \{k\}$). Thus the projection $Z_n = \pi(Y_n)$ on $(0,1]$ converges in distribution to the projection $\pi(Y)$, which is a compound Poisson process with compounds identical to k.

Remark 3.4: The particular case when $D(u_n)$ and $D'(u_n)$ hold, is included. For $D'(u_n)$ implies $\alpha_2=0$, thus $s_0=2$, $E_n^{(1)} \to \tau > 0$ and $\mu_1 = \alpha_1/\alpha_1 = 1$, $\mu_k = 0$ for $k \geq 1$. Thus (A) and (4), therefore also (A) and (6), are satisfied.

Remark 3.5: Since (3) and (4) imply (A) by Lemma 2.1, the theorem is also true if (A) and (6) is replaced by (3) and (4).

For general situations of the mark measure μ, we need a stronger condition D* instead of D.

<u>Condition $D^*(u_n)$:</u> We assume that for any integers n,ℓ

$$\sup_{A,B} | P(A \cap B) - P(A)\cdot P(B) | \leq \alpha^*_{n,\ell} \text{ and } \alpha^*_{n,\ell_n} \to 0 \text{ as } n \to \infty,$$

where $A \in \sigma\{U_1, U_2, \ldots, U_k\}$, $B \in \sigma\{U_{k+\ell+1}, U_{k+\ell+2}, \ldots, _n\}$ for any k with $U_i = 1\{X_i > u_n\}$ and a suitable sequence $\ell_n = o(n)$.

Note that $D^*(u_n)$ is still weaker than the strong-mixing condition.

Let the measure $\lambda\times\mu$ be defined by $\lambda\times\mu(B) = \theta\tau\sum_j (b_j - a_j)\mu(N_j)$ for B as in (8).

<u>Theorem 3.6</u>: Assume that the stationary sequence $\{X_n\}$ satisfies $D^*(u_n)$, (A) and (6). Then $Y_n \overset{d}{\to} Y$ where Y is a Poisson process on $(0,1] \times \mathbb{N}$ with intensity measure $\lambda\times\mu$. Thus the projection $Z_n = \pi(Y_n) \overset{d}{\to} \pi(Y) = Z$ with Z a compound Poisson process where the probability law of the compounds is given by μ.

<u>Proof:</u> By Kallenberg's Theorem and Lemmas 2.3 and 3.1, it is sufficient to prove $P\{Y_n'(B) = 0\} \to \exp(-\theta\tau \sum_{j=1}^{J} (b_j - a_j)\mu(N_j))$ with B as in (8). Now analogously to the proof of Theorem 3.3

$$P\{Y_n'(B) = 0\} = P\{\bigcap_{j=1}^{J} (N_n'(i) \notin N_j,\ i \in (k_n a_j, k_n b_j])\} =$$

$$= \prod_{j=1}^{J} \prod_{i\in(k_n a_j, k_n b_j]} P\{A_i\} + \sum_{\ell=2}^{|I_n|} \{P(\bigcap_{k=1}^{\ell} A_{i_k}) - P(\bigcap_{k=1}^{\ell-1} A_{i_k})\cdot P(A_{i_\ell})\}$$

where $A_i = \{N_n'(i) \notin N_j\}$ with $j = j(i)$ and by enumerating the i's in $I_n = \bigcup_{j=1}^{J} (k_n a_j, k_n b_j]$ with i_k.

Each term of the sum is by $D^*(u_n)$ bounded by α^*_{n,ℓ_n} since the index sets are separated by ℓ_n. Thus the sum is bounded by

$$O(|I_n|\alpha^*_{n,\ell_n}) = O(k_n \cdot \alpha^*_{n,\ell_n}) = o(1)$$

by choosing k_n such that $k_n \alpha^*_{n,\ell_n} = o(1)$ and $k_n \ell_n = o(n)$

(e.g. $k_n = \min(\alpha^{*-\frac{1}{2}}_{n,\ell_n}, (n/\ell_n)^{\frac{1}{2}})$. Finally, the product is equal to

$$\prod_{j=1}^{J} [P\{N_n'(1) \notin N_j\}]^{k_n(b_j-a_j)} = \prod_{j=1}^{J} \left(1 - \frac{\theta\tau\mu(N_j) + o(1)}{k_n} \right)^{k_n(b_j-a_j)}$$

$$\to \exp\left(-\theta\tau \sum_{j=1}^{J} (b_j-a_j)\,\mu(N_j)\right) = \exp(-\lambda\times\mu\,(B)) \text{ by using (6).}$$

Corollary 3.7: The statement of Theorem 3.6 is true if (3) and (4) hold together with $D^*(u_n)$.

4. Examples

In this section we discuss two examples exhibiting the particular cases given in Theorem 3.6 with $s_0 < \infty$ and $s_0 = \infty$.

1) An example, given in Haiman (1981), illustrates also Corollary 3.7 with $s_0 < \infty$. Let $\{\eta_k\}_{k\geq 0}$ be an iid. sequence with continuous distribution function $F(x)$. Let $\{J_k\}_{k\geq 1}$ be another iid. sequence, independent of $\{\eta_k\}$, with J_k Bernoulli (p), i.e.

$$0 < P\{J_k = 0\} = q = 1-p = 1-P\{J_k = 1\} < 1.$$

Then define $X_k = \eta_{k-J_k}$. Obviously, $\{X_k\}_{k\geq 1}$ is strongly stationary with marginal distribution $F(x)$. Let u_n be such that $n\overline{F}(u_n) = \tau > 0$. Note that $\{X_k\}$ is 2-dependent, thus $D(u_n)$ and $D^*(u_n)$ hold with any $k_n = o(n)$. We show now that (3) and (4) are satisfied. For

$$E_n^{(2)} = k_n \sum_{j=2}^{r_n} (r_n-j)\, P\{X_1 > u_n, X_j > u_n\} = k_n(r_n-1)\, P\{X_1 > u_n, X_2 > u_n\}$$

$$+\, O(k_n r_n^2\, \overline{F}^2(u_n)) = q\cdot p\, k_n(r_n-1)\overline{F}(u_n) + O(k_n^{-1}) \to \tau q\, p \text{ , since } k_n r_n \sim n$$

and $P\{X_1 > u_n, X_2 > u_n\} = qp\,\overline{F}(u_n) + (1 - q\,p)\,\overline{F}^2(u_n)$.

Furthermore we find that $E_n^{(3)}$ is bounded by $O(k_n r_n^2 \overline{F}^2(u_n)) = o(1)$.

Thus $\alpha_1 = \tau$, $\alpha_2 = \tau q p$, $\alpha_3 = 0$, $s_0 = 3$. Thus (3) and (4) hold with $\theta = \tau^{-1}(\tau - \tau q p) = 1 - qp > 0$. Finally

$$\mu_1 = (1-2\cdot qp)/(1-qp) \text{ and } \mu_2 = qp/(1-qp).$$

Thus Corollary 3.7 implies that the number of exceedances of the level u_n in (0,1] is asymptotically compound Poisson with mean number of clusters equal to $\theta\tau$ where the size of the clusters is either 1 or 2 with the above asymptotic probabilities.

2) The second example exhibits the case $s_0 = \infty$. We use the example of Denzel and O'Brien (1975) of a "chain-dependent" sequence $\{X_k\}_{k\geq 1}$ defined by means of an ergodic Markov chain $\{J_k, k\geq 0\}$ with positive integers as states and connected by

$$P\{J_n = j, X_n \leq x | X_1,\ldots,X_{n-1}, J_0, J_1,\ldots,J_{n-2}, J_{n-1} = i\}$$
$$= P\{J_n = j, X_n \leq x | J_{n-1} = i\} = P_{ij}H_i(x), \ \forall n\geq 1, \ i,j\geq 1.$$

P_{ij} are the transition probabilities, $P_{ij} = \theta\Pi_j + (1-\theta)\delta_{ij}$, with $\theta \in (0,1)$, $\Pi_j = j^{-1/2}-(j+1)^{-1/2}$, $j\geq 1$, $\Pi_0 = 1-2^{-1/2}$, $\delta_{ij} = 1$ for i=j, 0, otherwise. $H_i(x)$ are non-degenerate distribution functions defined in the following. Let $H(x)$ be a continuous distribution, $\overline{H}(x) = 1-H(x)$ and y_i such that $y_0 = -\infty$, $H(y_1) = \Pi_1$, $H(y_i) = \Pi_1 + \Pi_2 + \ldots + \Pi_i = 1-(i+1)^{-1/2}$. Then let

$$H_i(x) = \begin{cases} 0 & \text{if } x\leq y_{i-1} \\ \Pi_i^{-1}(H(x) - H(y_{i-1})) & \text{if } y_{i-1} \leq x \leq y_i \\ 1 & \text{if } x > y_i \end{cases}$$

In the stationary case (i.e. the distribution of J_0 is $\Pi = (\Pi_0, \Pi_1, \ldots)$), X_n has the marginal distribution $H(x)$. Since this "chain dependent" sequence $\{X_n\}$ is strong-mixing, the condition $D^*(u_n)$ is satisfied with u_n such that $n\overline{H}(u_n) = \tau > 0$. A simple argument shows that

$$y_{j(n)-1} = d_n^* \leq u_n(\tau) \leq d_n = y_{j(n)}$$

with $j(n) = [n^2/\tau^2]$ and $n\overline{H}(d_n) \to \tau$, $n\overline{H}(d_n^*) \to \tau$. Therefore we may consider exceedances $N_n(i)$ of the level d_n, which simplifies the calculations. We have to prove that $P\{N_n(1) \geq k | N_n(1) \geq 1\} \to \sum_{\ell=k}^{\infty} \mu_\ell = (1-\theta)^{k-1}$, $\forall\ k \geq 1$. Thus μ is the geometric distribution on $\mathbb{N}$.

i) Note that following Denzel and O'Brien (1975)

$$P\{N_n(1) = 0\} = P\{M_{r_n} \leq y_{j(n)}\} = \{1-(1+j(n))^{-\frac{1}{2}}\}\{1 - \theta(1+j(n))^{-\frac{1}{2}}\}^{r_n-1}$$

and thus

$$P\{N_n(1) \geq 1\} = (1+o(1))\ \theta\ r_n\overline{H}(d_n).$$

ii) We deal now with $P\{N_n(1) \geq k\}$. We use that by the construction of the sequence X_n and d_n

$$\{N_n(1) \geq k\} = \{\#\ \{i: J_i > j(n),\ i=0,\ldots,r_n-1\} \geq k\} = A_k$$

For k fixed we denote the times of the first k exceedances $J_i > j_n$ by $i_1, i_2, \ldots, i_k$ with $0\ i_1 \leq i_2 \leq i_3 \leq \ldots \leq i_k \leq r_n-1$.

$$B = \{\forall\ \ell = 1,\ldots,k-1:\ \ i_{\ell+1} = i_\ell + 1\}.$$

In the following we consider the events $A_k \cap B$ and $A_k \cap B^c$, and use the following transition probabilities

a) $p_{ih}^{(k)} = P\{J_{k+1} = h | J_1 = i\} = \Pi_h(1-(1-\theta)^k) + (1-\theta)^k\ \delta_{ih}$

$\forall\ i \geq 1,\ h \geq 1,\ k \geq 1.$

b) $p_{ih}^{*(k)} = P\{J_{k+1} = h,\ J_k \leq j(n),\ldots,J_2 \leq j(n) | J_1 = i\} =$

$$= \begin{cases} \theta^2 H(d_n)\Pi_h(1-\theta\overline{H}(d_n))^{k-2} & \text{if } k>1 \\ \theta\Pi_h + (1-\theta)\delta_{ih} & \text{if } k=1 \end{cases} \qquad i,h>j(n).$$

c) $P^{*(k)}_{ih} = \theta\Pi_h\ (1 - \theta\overline{H}(d_n))^{k-1}$ if $k\geq 1$, $i\leq j(n)$, $h>j(n)$.

By using these formulas, a straightforward calculation shows that

$$P(A_k \cap B) \sim (1-\theta)^{k-1}\ \overline{H}(d_n)\ (1 + \theta H(d_n) \sum_{i=1}^{r_n-k+1} (1-\theta\overline{H}(d_n))^{i-2}).$$

Since $(1 - \theta\overline{H}(d_n))^{r_n-k-1} = 1 - \theta r_n\overline{H}(d_n) + o(r_n\overline{H}(d_n))$, we find that

$P(A_k \cap B)/P\{N_n(1) \geq 1\} \to (1-\theta)^{k-1}$, $k\geq 1$.

Using the same formulas for $P(A_k \cap B^c)$ we find that $P(A_k \cap B^c) = o(r_n\overline{H}(d_n)) = o(P\{N_n(1) \geq 1\})$, which completes the proof on the cluster size distribution μ.

Therefore Theorem 3.6 implies that the point process Y_n converges to the point process Y in distribution with

$$\mu_k = \theta\cdot(1-\theta)^{k-1}, \qquad k\geq 1.$$

In the case $\theta > \frac{1}{2}$ it is also possible to prove the conditions (3) and (4). But for $\theta < \frac{1}{2}$ condition (3) is not satisfied, thus showing that (3) is not a necessary condition.

From the definition of B and the above derivation it follows also that if the sequence X_k exceeds the level u_n, then this happens consecutively a geometric random number of times.

References

Denzel, G.E. and O'Brien, G.L. (1975) "Limit theorems for extreme values of chain-dependent processes. Ann. Probab. 3, 773-779.

Feller, W. (1968) "An introduction to probability theory and its applications". Vol. I, 3 ed. Wiley, New York.

Haiman, G. (1981) "Valeurs extrémales de suites stationnaires de variables aleatoires m-dépendantes". Ann. Inst. Henry Poincaré, 17, B, 309-330.

Kallenberg, O. (1976) "Random measures." Academic Press.

Leadbetter, M.R. (1983) "Extremes and local dependence in stationary sequences". Z. Wahrscheinlichkeitstheorie 65, 291-306.

Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983) "Extremes and related properties of random sequences and processes." Springer Statistics Series. Springer, Berlin.